

A first exploration of word- and phrase-final f0 movements in spontaneous Yali

Constantijn Kaland, Katharina Gayler, Martine Grice, Nikolaus P. Himmelmann

Article - Version of Record

Suggested Citation:

Kaland, C., Gayler, K., Grice, M., & Himmelmann, N. P. (2026). A first exploration of word- and phrase-final f0 movements in spontaneous Yali. *Phonological Data & Analysis / Linguistic Society of America*, 8, Article e4. <https://doi.org/10.1017/s2642182825000030>

Wissen, wo das Wissen ist.

This version is available at:

URN: <https://nbn-resolving.org/urn:nbn:de:hbz:061-20260707-114711-5>

Terms of Use:

This work is licensed under the Creative Commons Attribution 4.0 International License.

For more information see: <https://creativecommons.org/licenses/by/4.0>

RESEARCH ARTICLE 

A first exploration of word- and phrase-final f_0 movements in spontaneous Yali

Constantijn Kaland¹, Katharina Gayler³, Martine Grice² and Nikolaus P. Himmelmann³

¹Linguistics I - Phonetics and Phonology, Heinrich Heine University Düsseldorf, Germany; ²Institute of Linguistics – Phonetics, Universität zu Köln, Germany and ³Institute of Linguistics, Universität zu Köln, Germany

Corresponding author: Constantijn Kaland; Email: kaland@hhu.de

(Received 22 March 2023; accepted 14 March 2024)


Abstract

The current study reports a first explorative analysis of f_0 movements in final and non-final positions in Yali, a Trans-New Guinea language spoken in the Papuan highlands in Indonesia. The language is understudied, in particular for its prosody. The available literature hints at the existence of word-level f_0 movements. However, acoustic analyses are lacking and it remains a challenge to disentangle word-level f_0 movements from those at the phrase level, due to their interaction. This study analyses spontaneously produced speech acoustically for f_0 in final and non-final syllables at the word and phrase level, and by means of a cluster analysis to reveal the most common f_0 patterns at either prosodic level. Results provide strong support for word-final rising f_0 movements that are particularly clearly realised in non-final words in the phrase. The outcomes are tentatively interpreted within their phonological context and in terms of the different levels that could be distinguished in Yali prosody.

Keywords: Yali; prosody; f_0 ; intonation; acoustic analysis; cluster analysis

1. Introduction

Prosody, that is, suprasegmental aspects of speech and language such as stress, rhythm and intonation, often remains understudied in language documentation (e.g. Himmelmann & Ladd 2008). In part, this is due to the complexity of the prosodic structure, often affected by linguistic and paralinguistic factors (e.g. Ladd 2008). Furthermore, initial language descriptions often lack in-depth acoustic analyses that provide quantitative support for (impressionistically) observed phenomena. The study of prosody in underresearched languages makes a particularly valuable contribution towards refining existing theory. That is, typological accounts are often shaped by work on well-studied languages and may therefore be biased. This study focuses on Yali, a Trans-New Guinea language of Indonesia that is documented to some extent, but lacks contemporary analysis of its prosodic structure (Riesberg et al. 2021). The limited studies available for Yali and other members of its language family suggest that Yali has a form of word prosody, possibly using word-final tones. The current study provides an overview of the available literature on Yali and related languages and then explores the nature of Yali's word-level prosody using cluster analysis on f_0 contours (Kaland 2021). This analysis offers a solid starting point for prosodic analysis due to its data-driven approach and reproducible nature. Although the current paper covers a specific feature of Yali prosody and leaves several aspects for future research, it offers a pioneering in-depth study of an otherwise unknown prosodic system using modern analysis methods and, crucially, analysing naturalistic spontaneously produced speech.

 This research article was awarded Open Materials badge for transparent practices.

© The Author(s), 2026. Published by Cambridge University Press on behalf of Linguistic Society of America. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

The following (sub)sections discuss three different typological accounts of prosodic prominence and the role of f_0 in languages with different prosodic structures (Section 1.1). Thereafter, the Yali language situation and the available studies on the prosody of related Dani languages are reviewed (Section 1.2). Note that the literature review concerns all relevant works known to the authors to date and can therefore be taken as a reference for future work on prosody of languages in this area. The research questions of this study are formulated in Section 1.3.

1.1. F_0 in prosodic typology

Ladd (2008:165) proposes an intonational typology along two dimensions (Table 1), one of which is based on Beckman 1986 (see Lindström & Remijsen 2005 for a similar proposal which provides a space for languages without any pitch accents such as Wolof). The typology distinguishes four categories of languages, based on whether phonetic cues to word stress are used to signal the phonologically strongest syllable in a word or not (stress-accent or non-stress accent in Beckman 1986) and whether they use f_0 only at the phrase level (post-lexical) or at both word and phrase level (lexical pitch). Swedish and Serbo-Croatian are both categorised as stress languages that make use of lexical f_0 contrasts. A common example from Swedish is *stegen* (e.g. Bruce 1977); either ['stê:gen] 'steps' (accent 1) or ['stë:gen] 'ladder' (accent 2). Note that in both words the first syllable is stressed. Japanese, a language without stress accent, shows f_0 contrasts; either an accent on the first syllable [kãmi] 'God' or on the second syllable [kàmí] 'hair'/'paper' (Beckman 1986). In English, the minimal stress pair *permit*, either ['pərmít] (noun) or [pər'mít] (verb) differs acoustically mainly in vowel quality, whereas f_0 does not play a consistent role in signaling word meaning distinctions. At the phrase level, however, English is claimed to have different pitch accent types that highlight important information (i.e. focus) or demarcate (intermediate) phrases (e.g. Pierrehumbert & Hirschberg 1990). Bengali is a language that has no lexical specification of pitch, no or weak phonetic marking of word stress, and only uses f_0 at the phrase level (focus marking, boundary tones; Hayes & Lahiri 1991; see also Ladd 2008).

In another account by Gordon (2014), word stress and pitch accents are disentangled in a typology where prosodically prominent syllables are found across languages (Table 2). A central assumption of this account is that word stress patterns reflect phrasal prominence, in particular due to **edge repulsion**.

Table 1. Prosodic typology in Ladd (2008).

		Phonetic typology	
		Stress accent	Non-stress accent
Lexical typology	Lexical pitch	example: Swedish	example: Japanese
	Post-lexical pitch only	example: English	example: Bengali

Table 2. Prosodic typology in Gordon 2014; PA = pitch accent.

	Edge repulsion		
	Word	Phrase	Languages
Symmetrical	Yes	Yes	English, Egyptian Arabic
	No	No	Hebrew, Farsi
Asymmetrical	No	Yes	Chickasaw, Cayuga, etc.
	No	No PA	Wolof
	Yes	No	<i>Unattested?</i>
	Yes	No PA	<i>Unattested?</i>

Edge repulsion is a compensatory mechanism to avoid tonal crowding at phrase edges. That is, if the realisation of a boundary tone plus another (phrasal) tone is phonologically required to appear at the right edge of a phrase, that is, on the final syllable, some languages choose to realise the phrasal tone on the pre-final syllable. For example, in American English a prenuclear H* pitch accent is realised earlier in the prenuclear syllable when immediately followed by the nuclear accented syllable (Silverman & Pierrehumbert 1987). In Papiamentu the LH tone is also realised more leftward (away from the boundary) when occurring phrase-finally (Remijsen & Van Heuven 2005). The typology predicts that in languages with non-peripheral (e.g. penultimate or antepenultimate) stress there must be edge repulsion at the phrase level (Gordon 2014). English and Egyptian Arabic are examples of such languages, in which word and phrase level are symmetrical in the sense that edge repulsion applies to both. Whereas in languages in which there is no edge repulsion at the phrase level, stress can be found on either a peripheral syllable (ultimate, e.g. Hebrew, symmetrical) or a non-peripheral one ((ante-)penultimate, e.g. Chickasaw, asymmetrical). A fourth category consists of asymmetrical languages that have word stress but do not make use of pitch accents that align with stressed syllables and therefore have no reason to apply edge repulsion (e.g. Wolof). Although hypothetically, languages might have edge repulsion only at the word level and not at the phrase level, these have so far not been attested. The unattested categories confirm the theory that if edge repulsion is found at the word level it must reflect the location of prominence at the phrase level, 'since there are less likely to be any pitch properties bounded by the word that would repel stress from an edge' (Gordon 2014:20).

The most recent overview of prosodic prominence across languages by Jun (2014) (following Jun 2005) proposes a typological account according to three major aspects: prominence type, word prosody, and macro-rhythm (Table 3). Prominence type refers to the way a language marks phrasal prominence by f0. This can be done by means of pitch accents on the lexically or post-lexically specified head of the phrase (**head languages**). Languages belong to this prominence type category when they only mark the head as prominent (not the edge). For example, English and Arabic (stress) or Swedish (stress and tone) are head languages. Other languages mark both the head and the edge of a prosodic domain; that is, word

Table 3. Prosodic typology in Jun (2014) Table 17.4. Shaded cells indicate language categories that are unexpected to be attested.

Prominence type	Word prosody	Macro-rhythm		
		Strong	Medium	Weak
Head	Stress	Brazilian Portuguese, Castilian Spanish, Catalan, Egyptian Arabic, Greek, Italian, Samoan	Dutch, English, German, Jamaican Creole, English, Lebanese Arabic	European Portuguese, Wolof, Kuot
	Tone/lexical pitch accent	Kihehe, Safwa, Somali	Isthmus Zapotec, Kimatuumbi, Skou	Cantonese
	Both	Papiamentu, Swedish	—	Mandarin
	None	—	—	
Head/edge	Stress	Bengali, Georgian, Kiche, Tamil, Tongan	Bininj Gun-wok, Dalabon	
	Tone/lexical pitch accent	Japanese, Leketio Basque	—	
	Both	Serbo-Croatian	Chickasaw	
	None	French, Standard Basque	—	
Edge	None	Accentless dialects of Japanese, Halh Mongolian, Oirat Mongolian, Seoul Korean, West Greenlandic		

or phrase. These **head/edge languages** can have a lexically specified head (e.g. Tokyo Japanese) or a post-lexical one (e.g. French). The edges in this type of languages are often marked at the level of the accentual phrase (AP; Pierrehumbert & Beckman 1988), a level corresponding to one or more words, approximating the phonological phrase level (ϕ in Nespor & Vogel 2007). **Edge languages** are claimed not to have any (post-)lexically specified head that is marked with prosodic prominence, but to only mark the edges of prosodic phrases (on a number of different levels). These languages are reported to not exhibit any word prosody (e.g. Seoul Korean).

As for the head and head/edge languages, different types of word prosody are distinguished. A language could have word stress (e.g. English and Bengali), some tonal marking at the lexical level (tone, e.g. Somali, or lexical pitch accent, e.g. Japanese) or both of these (e.g. Swedish and Serbo-Croatian). The typology in Jun 2014 does not exclude head languages for which the head is only post-lexically specified (i.e. no word prosody), however, these head languages are unattested. Post-lexical marking of prosodic heads is attested in two head/edge languages, albeit with a strong tendency to occur at the right word edge (French and Basque). Thus, the typology strongly suggests that prominence marking in non-edge phrase positions (i.e. phrase medially) is mainly driven by word prosody, similar to Gordon's typology (Gordon 2014). In this respect there is an important trade-off between head or edge marking as a means to make the word prominent: '... edge tones will help to mark the prominence of [a] word when the head marking (stress) is weak, which is common in head/edge-prominence languages' (Jun 2014:537).

The concept of **macro-rhythm** was introduced to capture similarities between languages that differ with respect to prominence type and/or word prosody (Jun 2014). Macro-rhythm concerns the regularity with which high and low tones alternate in a phrasal intonation contour (as opposed to micro-rhythm, which concerns the smaller units such as syllables or feet). Three degrees of macro-rhythm are distinguished; strong, medium, and weak. This degree is determined by (i) the number of possible phrase-medial pitch accents with larger pitch accent inventories leading to weaker degrees of macro-rhythm at the phrase level, (ii) the type of the most common phrase-medial pitch accent with rising or falling accents being more rhythmic than level ones, and (iii) the frequency or domain of pitch accents, with smaller domains (e.g. the word) being often marked, leading to a higher degree of macro-rhythm. As a consequence, the less marking of heads the stronger the degree of macro-rhythm, to the extent that all edge languages are classified with a strong degree of macro-rhythm (Table 3). It is important to note that the concept of macro-rhythm constitutes an overall tonal regularity that surfaces within a phrase. That is, phonological rules governing the association of pitch accents to stressed syllables are not immediately relevant for determining the degree of macro-rhythm, because a subunit of tonal rhythm (e.g. LH in L-H-L-H-L-H) 'can cover materials larger or smaller than a word ... and does not necessarily form a prosodic unit' (Jun 2014:524).

The typological accounts of Ladd (2008), Gordon (2014), and Jun (2014) have two important aspects in common. First, they both allow for a special status of f_0 compared to other acoustic cues such as duration or intensity. This special status is reflected in the central role of specific tonal phenomena such as lexical tone contrasts (Ladd 2008), edge repulsion (Gordon 2014), or tonal rhythm (Jun 2014). Second, f_0 plays the most important role at the phrase level and a secondary role at the word level. That is, languages are assumed to show at least post-lexical use of f_0 (Ladd 2008) and to exhibit some degree of tonal macro-rhythm, irrespective of whether a language has word prosody (Jun 2014). And importantly, edge repulsion at the phrase level may determine the location of prominences at the word level, but not vice versa (Gordon 2014). The latter point requires careful and systematic analyses to reveal the potential interplay of word-level and phrase-level prosody, in particular in languages that make use of tonal targets at both levels. Yali may be such a language and is further discussed in the following section.

1.2. Language situation

Yali is spoken in the area east of Wamena, the capital of the Jayawijaya district in the Papuan Highlands (Papua Pergunungan, Indonesia). It is a member of the Dani family, spoken in the Baliem Valley and surrounding highlands. The language is 'stable' (Eberhard *et al.* 2020) and between 4 and 5 on the EGIDS

scale (Lewis & Simons 2010, Anderbeck 2015). Classification of these languages and of Papuan languages in general is subject to ongoing revision given that they are often underresearched (e.g. Palmer 2018, Usher & Suter 2021). The name **Yali** (iso: yac) refers to a Trans-New Guinea dialect cluster with a total number of speakers estimated at around 30,000, according to several sources (Silzer & Heikkinen-Clouse 1991, Wilson 1999, Lewis et al. 2015, Riesberg et al. 2021). However, these counts all lead to the same original sources (Voorhoeve 1975 and Grimes 1984), the validity of which to date could be questioned. It has been suggested that there are three Yalic languages (Hammarström et al. 2021): Angguruk, Ninia, and Pass Valley Yali. However, it remains unclear how many dialects can be subsumed under Pass Valley Yali. Hammarström et al. (2021) report three dialects: Apahapsili, Landikma, and Pass Valley Yali. Fahner (1979) and Riesberg (2021) also list Abenaho as an additional dialect.

The available published linguistic work on Yali is limited. Most recently, Angguruk Yali was covered in a dictionary (Riesberg et al. 2021), which includes a grammar sketch that mainly focuses on the verbal system. Abenaho Yali morphology is studied in Fahner (1979). A small number of unpublished master's theses have focused on specific aspects of Yali; the verbal system (Sawaki 1997), affixation (Walianggen 2012a), verb agreement (Walianggen 2012b), causative constructions (Lekeneny 2014), and reduplication (Wandik 2014). As for phonology, existing work does not go beyond sound inventories. Angguruk Yali is reported to have the following vowels: /i, i, u, u, e, o, a/ and consonants: /p, t, k, φ, s, χ, m, n, ŋ, l, w, j/ (Riesberg 2021). Thus, prosodic aspects of Yali remain to be studied. Note, however, that more elaborate phonological descriptions do exist for other Dani languages found in areas bordering on the ones where Yali is spoken. It is relevant to also consider these, as neighboring language varieties have been reported to be mutually intelligible in this area (e.g. Heeschen 1998). This will be done in the following discussion, which is limited to languages in the Dani family as classified in Glottolog (Hammarström et al. 2021).

1.2.1. Lower Grand Valley Dani

The phonology of Lower Grand Valley Dani (iso: dni), spoken south of Wamena, has been covered extensively in Bromley 1961. The analysis is couched in the Pikean version of the American structuralist model of intonation (Pike 1945), which distinguishes (primary) stress from four contrastive pitch levels associated with the boundaries of intonationally defined units on a number of different hierarchical levels. This model has been criticized in a number of regards (cp., for example, Lieberman 1965, Crompton 1980, Ladd 2008:62–65).

One problem pertains to the fact that analyses presented in this framework are difficult to interpret without access to a recording. Bromley's analysis of Dani—one of the most detailed analyses of the prosodic system of a Papuan language published in the twentieth century—is, unfortunately, no exception. Here we only attempt to summarize the points that are most salient to us and provide a basis for generating testable hypotheses. In Bromley's analysis, the phonological word in Dani is defined as the unit carrying a single primary stress. This word stress occurs on the final syllable of the stem. The stem may be followed by one or more clitics, which are unstressed. The acoustic cues for word stress are 'upgliding pitch and/or increased loudness' (Bromley 1961:49). The details of the pitch glide 'are matters of intonation' (loc. cit.) and appear to be influenced primarily by position in the phrase, but also by word length.

The next level in the hierarchy of prosodic units is the 'intonation contour group', which consists of one or more words (most examples provided by Bromley only encompass one word). It is at this level that contrastive pitch levels, which are combined to describe distinctive tunes, are assigned (Bromley 1961:53). The intonation contour group is further defined by including one syllable carrying primary contour stress, which involves lengthening of the syllable nucleus in addition to being associated with a contrastive pitch level. The stressed syllable is generally the final syllable of the stem of its last word.

The third level in Bromley's prosodic hierarchy is called the 'phonological sentence', which consists of one or more intonation contour groups. This level is essentially defined by the 'final relaxation of the articulatory muscles' (Bromley 1961:58). Otherwise, phonological sentences involve relatively minor modifications of its final intonation contour group.

A major shortcoming of Bromley's analysis is that it does not propose a clear inventory of possible tunes. That is, in principle it appears to be possible that the four pitch levels all can occur on the contour

group-final stressed syllable and that they may be combined rather freely with each other (i.e. level 1 may follow level 3 and vice versa, level 4 level 2, etc.). Another problem of interpretation pertains to the question of how Bromley's prosodic levels relate to the levels of the prosodic hierarchy in Nespor & Vogel 2007. The correspondence is unproblematic in the case of the phonological word as this unit appears to be the same in both hierarchies. The 'intonation contour group' appears to cover both the intonational phrase (IP) and the intermediate (or phonological) phrase (ip), as it involves boundary tone phenomena and allows for pauses to occur. For current purposes, we simply equate it with the IP, as the distinction between these two levels in the Nespor & Vogel hierarchy is not clear in all regards (see Ladd 2008:274–84, 297–308 for pertinent discussion). The 'phonological sentence', finally, appears to pertain to a level higher than the IP for which there is very little terminological consensus in current work, 'utterance' perhaps being the most widely used term. Despite its shortcomings, Bromley's analysis constitutes a number of useful observations and testable claims that warrant further investigation. Most important for the current study is the analysis of Lower Grand Valley Dani as having word-level stress that is signalled by f_0 .

1.2.2. Wano

Spoken north-west of Wamena, Wano (iso: wno) is another Dani language, the phonology of which has been studied to some extent (Burung 2007, 2017). These works claim regular ultimate word stress, also for affixed words, but not for cliticized words, as clitics remain unstressed (see 1, taken from Burung 2017:64). No information is provided about how stress is realized.

- (1) a. suffix {-mu} *ultimate stress remains*
 pavuk [pa'βuk] 'bush'
 pavuk-mu [paβuk'mu] 'in (the) bush'
 b. clitic {=a} *ultimate stress becomes penultimate*
 maduk [ma'duk] 'bad'
 maduk=a [ma'duga] '(is it) bad?'

For phrase intonation, Burung (2017) distinguishes between falling, flat (most common), and rising. Falling contours are found at the end of an utterance, in questions without a compulsory answer, and in tail-head linked clauses, often accompanied by the clitic {=o} (pause). Flat contours accompany clauses that are chained. Rising contours are found in questions requiring a compulsory answer, for example with the clitic {=a} (question), and in simple (i.e. short) statements where the rise coincides with the stressed syllable.

1.2.3. Western Dani

An extensive grammar of Western Dani (iso: dnw) covers many aspects of this language (Barclay 2008), but fails to cover prosody. Although minimal, a phonological description of Western Dani does make reference to word prosody (Purba *et al.* 1993:36). A tentative analysis of ultimate word stress in three-syllable words (the most common word length reported) is given (see examples in 2) and it is concluded on the basis of a lack of minimal pairs that stress is not lexical.

- (2) *wumukor* [wumu'kor] 'moss'
 iniyok [ini'yo?] 'foot'
 lebenak [lebe'na?] 'weak'
 ogobak [ogo'ba?] 'nose'

1.2.4. Nggem

Nggem (iso: nbq) is spoken north-west of the Pass Valley Yali dialects. Its morphology and syntax is described in Etherington 2002. This language is claimed to have regular penultimate word stress, and although there are no minimal stress pairs, some exceptions in which stress is not penultimate are reported (Etherington 2002); see 3. No descriptions of f_0 contours are reported.

- (3) *abi* ['abi] 'nose'
abis [a'bis] 'mouth'
avula [a'pula] 'pregnant'
avulak ['apuluk] 'clatter'

1.2.5. Yali

To date, there is no literature on the prosody of Yali (iso: yac), although data collections are publicly available for the Angguruk and Apahapsili dialects (Riesberg et al. 2012). The Angguruk dictionary (Riesberg et al. 2021) as well as the other sources on Dani languages make no reference to lexical items that can be distinguished solely by f₀, which for now we take as a reason to exclude the option of classifying Yali as a language with lexically contrastive tone (see further discussion of lexical pitch accent languages in the summary sub-section below). From the available data (descriptions in Section 2) some examples are given here to illustrate the f₀ contours found in words and phrases.

At the word-level, final syllables are regularly found with a high or rising f₀ movement, whereas non-final syllables show lower-level movements; see the syllable *ho* in 4, and Figure 1. Note the transcription with initial 'h' (Riesberg 2021), which correspond to a voiceless [h] or voiced glottal fricative [ɦ] as an allophone of /ɣ/. The high movement is still found on the final syllable of the word if that word contains a clitic (final syllable *sat* and *ma* in 5 and Figure 2). Note that the classification of affixes vs. clitics given here is based on morphosyntactic behaviour. Both suffixes and enclitics are treated as word-final syllables for the purposes of placing the word-final rise: there are usually rises on suffixed or cliticized syllables and there never are rises on stem-final syllables when followed by a suffix or a clitic. In this regard, suffixed/

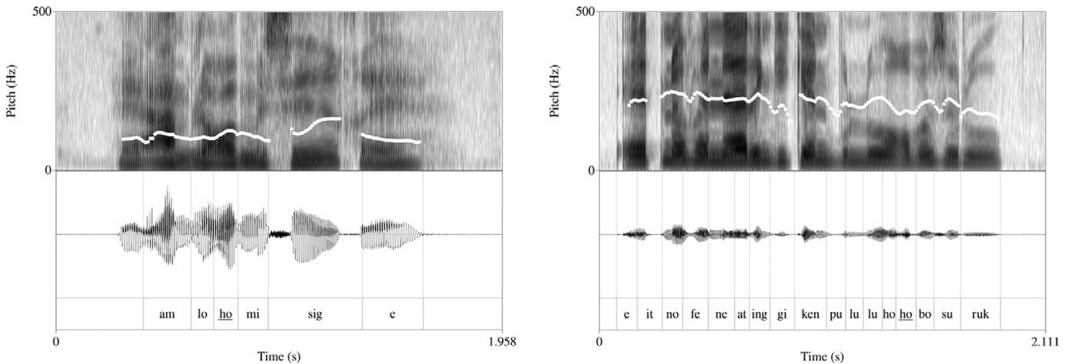


Figure 1. The syllable *ho* in word-final (left) and word-non-final (right) context (gloss in 4).

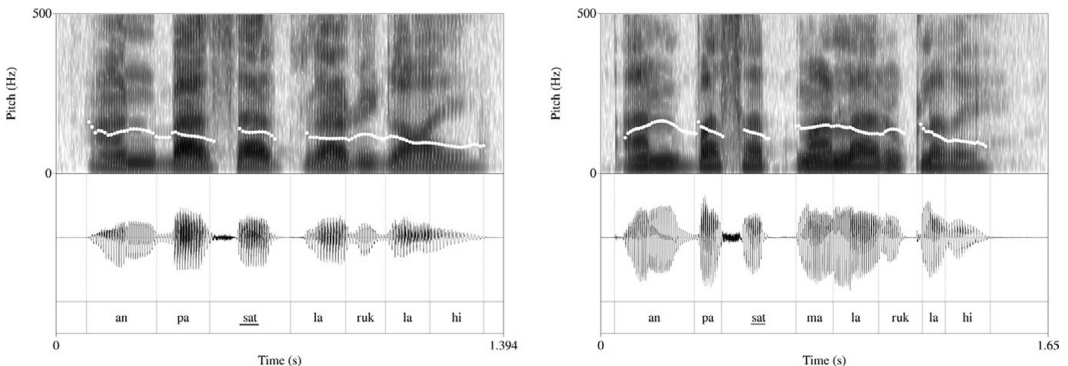


Figure 2. The syllable *sat* in word-final (left) and word-non-final (right) context (gloss in 5).

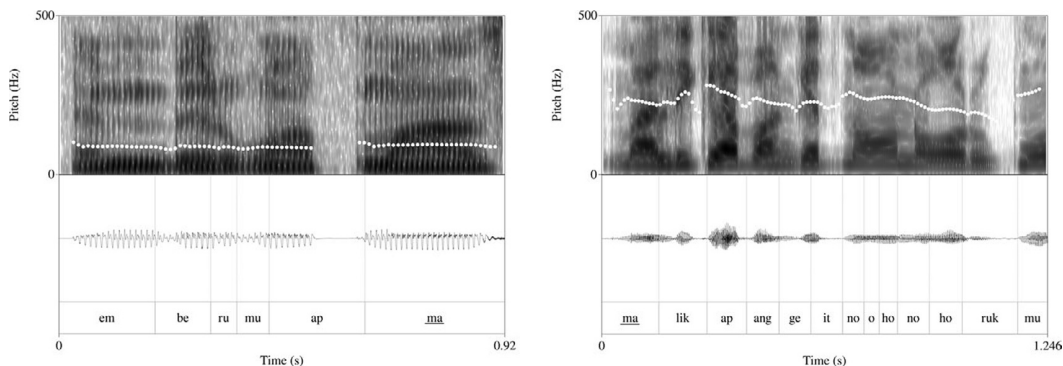


Figure 3. The syllable *ma* in word-/phrase-final (left) and word-/phrase-non-final (right) context (gloss in 6).

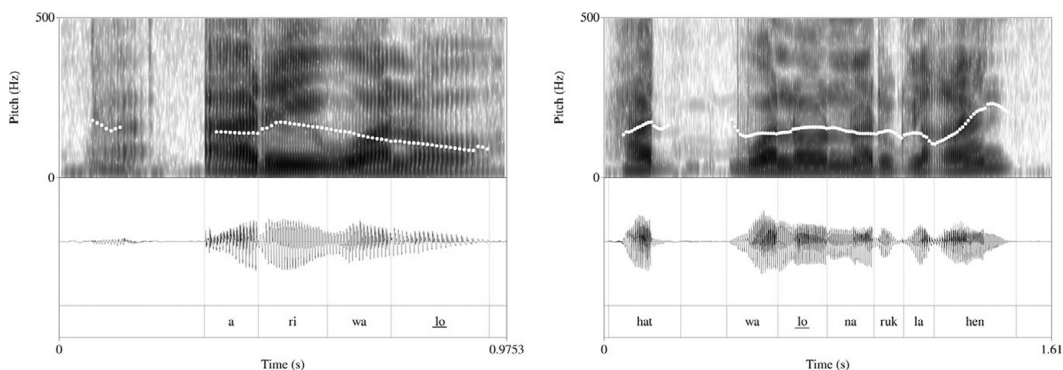


Figure 4. The syllable *lo* in phrase-final (left) and phrase-non-final (right) context (both cases word-final; gloss in 7).

enclitic syllables differ clearly from free/non-clitic grammatical function words where the final syllable of the preceding lexical item can carry a rising tone (e.g. *fam* ‘on/in’). This differs from the account given for Wano but it is very well possible that the author of the Wano description uses terminology differently. If the syllable is both word-final and phrase-final, the f_0 movement is not necessarily high or rising. Consider *ma* in 6 and Figure 3; the final movement does not clearly stand out as higher than the non-final syllable *ap*, but stays level. In 7 and Figure 4 the word- and phrase-final syllable *lo* has a falling f_0 movement, whereas in non-final phrase position the same word-final syllable is high.

It seems, therefore, that f_0 might play a role at the word-level in Yali prosody such that the final syllables of words have a high tonal target. This target is reached most clearly in non-final phrase positions by a high or rising f_0 movement. In phrase-final positions, the f_0 movement appears to be affected primarily by the demands of the (pragmatic) discourse context, clause type, and speech act. Falling movements are found in elicited statements (see 5 and 7a), rising movements accompany elicited questions (7b) and level or rising movements are found in spontaneous speech to indicate continuation (6).

(4) Word-level: final (a) and non-final (b) syllable ho

- a. *amloho misig e*
- amloho misig e*
- son one INTJ
- ‘his son’

- b. *e itno fene at inggiken pululu hohobo suruk*
e itno fene at inggik=en pululu hohobo su-tuk
 tree DET at just claw=CAUSE sound scratch do-PROG
 'he is still scratching that tree'

(5) Word level: final (a) and non-final (b) syllable sat

- a. *an pasat laruk lahi*
an pasat la-tuk la-ih
 1SG market go-PROG do-1SG.IM.PST
 'I am going to the market'
- b. *an pasatma laruk lahi*
an pasat=mu la-tuk la-ih
 1SG market=LOC go-PROG do-1SG.IM.PST
 'I am going to the market'

(6) Final (a) and non-final (b) syllable ma

- a. *emberumu apma*
emberu=mu apma
 on.top=LOC below
 'on the upper and on the lower side'
- b. *malik ap angge itno oho nohorukmu*
malik ap angge itno oho noho-tuk=mu
 child man CLASS DET also sleep-PROG=DS.SEQ
 'the boy is sleeping and'

(7) Phrase-level: final and non-final syllable lo

- a. *ari walo*
ari walo
 DEM.s snake
 'This is a snake'
- b. *hat walo naruk lahen*
hat walo na-tuk la-ehen
 2SG snake eat-PROG do-2SG.IM.PST
 'Have you eaten snake?'

1.2.6. Dani languages: summary

Although the literature on Dani languages is far from comprehensive in its analysis of prosody, the occurrence of ultimate word stress is reported for all languages for which a (minimal) phonological description has been published (dni, wno, dnw, and nbq). Furthermore, except for Nggem, all these sources mention at least one interaction between word-level and phrase-level prosody. F0 is a particularly relevant cue for this interaction (Bromley 1961), although it remains unclear what exactly the status of f0 is at either level. As for Yali, the few examples discussed above also hint at a word-final tonal target, although lexical contrasts based on f0 differences are unattested. Given that the word-final syllable appears to be the locus of the relevant f0 movement, it remains to be seen whether it could be classified as a **lexical pitch accent** (e.g. Beckman 1986, Hyman 2006, Ladd 2008:164–68). Note that lexical pitch accent languages are claimed to differ from tone languages in that the former apply a single word-level tone to one syllable in the word and are more prone to phrase-level intonation influences, whereas the latter apply them on multiple syllables in a more stable way (e.g. Ota et al. 2018). It is worth emphasizing that the notion of lexical pitch accent is controversial. It has been argued that lexical pitch accents represent just one way tonal distinctions may be organized in human languages (Hyman 2006). That is, so-called lexical pitch accent languages are just one type of (lexical) tone languages, other major types

being syllable tone languages and word tone languages. On the other hand, ‘accent’ has been argued to be an abstract prosodic category that differs fundamentally from (lexical) tone and allows for both tonal and non-tonal marking (Beckman 1986, Beckman & Venditti 2010, van der Hulst 2012). If Yali indeed has tones on word-final syllables, it remains to be investigated at which phonological level exactly these apply and how the interaction with phrase intonation can be characterised. Even for Swedish there are competing analyses, namely **binary** (Bruce 1977) and **privative** (Riad 2003), in which the f_0 movement is either present or absent rather than chosen from a set of possible alternatives. Also recall the discussion of Japanese (Section 1.1) in this respect. It would take more than a single study to tackle all the issues just discussed. The exploratory nature and goals of the current study is further described in the next section.

1.3. Research questions

The aims of this study can therefore be summed up by the following two research questions:

- I. Do specific f_0 movements align with word-final syllables (compared to non-final ones)?
- II. To what extent do word-level f_0 movements interact with phrase-level ones?

In the following, the research questions of the current study are further explained. Note, however, that the state of the research on Dani languages, in particular on Yali prosody, leaves many questions unanswered. Previous work on Dani languages has described a word-final prominence that is most likely cued by pitch. The central question in this study is therefore whether we can find acoustic evidence for word-level tonal targets. At the same time it is expected for f_0 to play a role at the phrase level, regardless of potential word-level tones. Thus, the second research question concerns the extent to which word-level f_0 movements interact with phrase-level tones. The minimal assumption regarding phrase prosody is that phrase-final syllables show a tonal movement that is different from other phrase positions. Which direction that phrase-final f_0 movement has (i.e. falling, level, or rising) requires detailed study of the context in which the phrase was uttered, taking into account pragmatic, information-structural and speaker-individual factors. Although it is plausible that, at the phrase level, Yali prosody shows similarities with Lower Grand Valley Dani (Bromley 1961), this cannot be taken for granted and will be investigated in the current study.

In the next sections we report an exploratory investigation of the f_0 movements at the word and phrase level, as found in spontaneously produced Yali speech.

2. Methodology

This section describes the collection, processing, acoustic measurements, and analysis of the Yali data.

2.1. Data collection

The Yali speech data investigated in this paper consists of mainly short stories and interactive dialogue tasks. They were collected in Apahapsili between 2014 and 2019 by Sonja Riesberg funded by the Volkswagen Foundation (DoBeS project ‘Documentation summits in the central mountains of Papua’) and the Australian Research Council. Among the collected materials are retellings of Pear stories (Chafe 1980), QUIS tasks (Skopeteas *et al.* 2006), Space games (Senft 2007), retellings of the Frog story (Mayer 1969), and Tangram tasks (as used in e.g. Savino *et al.* 2019). The recordings lasted from one and a half to seventeen minutes ($\mu = 4.6$ minutes) and were made using different devices (Sony HDR-PJ260, Sony HDR-CX250, Zoom Q8, or Zoom H4) in the houses of some of the speakers. A subset of the recordings used in this study were archived (see documentation information in Riesberg *et al.* 2012). In total, thirty-seven recordings from thirty different native Yali speakers were used. Their personal data could not always be accurately

documented. From twenty-seven speakers the age and gender was known: sixteen males and eleven females, age range: thirteen to fifty-three ($\mu = 28.7$ years). Informed consent was given by all participants.

2.2. Data processing

The recordings were transcribed into (intonation) phrases by two language experts with the help of native speakers. The segmentation was carried out at the level of intonation units (Chafe 1994), corresponding to intonation phrases in the phonological hierarchy (e.g. Nespor & Vogel 2007). The intonation units were defined on the basis of auditory boundary cues such as pitch modulations and pauses. In a previous crosslinguistic comparative study using Yali among other languages, these boundaries were found to be perceived in highly similar ways by native and non-native speakers (Himmelman et al. 2018). Based on the transcriptions the phrases were automatically split into words (based on spaces) and syllables (four types of syllables were distinguished: V, VC, CV, and CVC). The positions of the syllables were then indicated by two numbers for each level. At the phrase level, the number of the word in the phrase and the total number of words in that phrase were counted. At the word level, the number of the syllable in the word and the total number of syllables in that word were indicated. In this way, final versus non-final syllables were distinguished on each level such that all syllables not occurring in a final position were marked as non-final. No further distinction was made as to which non-final position in the phrase or word these syllables occupied.

Given the use of Indonesian loanwords in most of the recordings, all words were looked up in a Toolbox (SIL International 2021) lexicon compiled as part of the language documentation (Section 2.1) and selected if they matched with an entry. Matches also included forms with bound morphemes (affixed entries and words including clitics), as preliminary observations suggest that these do not necessarily affect final f_0 movements (see example 5, and discussion in Section 4). The selection was further limited to content words (verbs and nouns) as these are crosslinguistically the most informative words and tend to attract prosodic phenomena (i.e. tone, word stress, and/or (post-lexical) pitch accents). Selected syllables were further limited to the ones with a CV structure, which represents the most common syllable structure in Yali. In order to disentangle the two prosodic levels final and non-final syllables at either the phrase (ϕ) or word (w) level were marked. This resulted in three finality categories (ϕw): phrase-final, word-final (ff); phrase-non-final, word-final (nf); and phrase-non-final, word-non-final (nn) (see Table 4). Note that phrase-final syllables are inherently word-final, such that the category phrase-final, word-non-final (fn) does not exist. The final selection of syllables were the ones that occurred in all three level/finality categories. For example, the syllable [ma] was only selected if found in all three categories. This resulted in a set of 381 syllables (127 triplets, twenty-seven unique syllable types), for which the boundaries were manually indicated in Praat textgrids. The syllables were taken from words with a length of two up to five syllables ($\mu = 2.47$), which occurred in phrases with a length of one up to fifteen words ($\mu = 4.47$).

2.3. Acoustic measurements

Five syllables appeared to be produced either too quietly, too briefly, or with background noise such that they needed to be excluded from acoustic analysis. The remaining syllables ($N = 376$, from twenty-nine different speakers; see the appendix for finality type per speaker) were acoustically analysed for f_0 and

Table 4. Three categories of syllables (marked with \surd) according to prosodic level (phrase ϕ / word w) and finality (final/non-final).

		w	
		Final	Non-final
ϕ	Final	\surd ff	×
	Non-final	\surd nf	\surd nn

duration. That is, f_0 was measured in a time-series manner, specifically, at twenty evenly distributed time points in the syllable. The f_0 measurements were taken in Hertz (Hz) using Praat (Boersma & Weenink 2019) using the autocorrelation method, and using an f_0 window of 50 to 500 Hz with otherwise default pitch settings. The measurements were also checked and corrected for octave-jumps and f_0 measurement tracking errors. Voiceless parts of the contour were interpolated linearly to obtain an uninterrupted series of f_0 values. Interpolation is also perceptually motivated, that is, listeners perceive intonational prominence largely in uninterrupted ways (Mixdorff & Niebuhr 2013). Although f_0 jumps were mainly absent in the data, it is still possible that f_0 perturbations due to the effect of consonants affected the extracted contours. These perturbations could affect the cluster analysis if they concern a significant part of the f_0 contour (e.g. onset level, register). This issue is further discussed in the final section.

Time-series measures furthermore allow us to compare contours of different lengths in a more equal way and provide a more precise measurement of the f_0 contour compared to static measures such as f_0 mean or f_0 range. F_0 measurements were speaker corrected, to account for effects of gender on the f_0 level. This was done using standardization, as proposed for word-level f_0 movements in Rose 1987. A crude categorical difference was made for each contour whether it was a rise or a fall. These categories were obtained by subtracting the standardised f_0 level of the last measurement point from the standardised f_0 level of the first measurement point. This resulted in an f_0 value indicating the direction and steepness of the f_0 contour (see Section 3 for the analysis of f_0 shape). In this way, negative values are obtained for falls and positive values are obtained for rises. The further the f_0 direction value is from zero, the steeper the f_0 movement is. Note that there were no cases in which multiple directions were found in one syllable (i.e. a rise+fall or a fall+rise). Duration was measured as the time (in ms) between the start and the end of the syllable. The measurements were obtained in Praat (Boersma & Weenink 2019).

2.4. Statistical analysis

Statistical analyses were carried out using R (R Core Team 2020, R Studio Team 2020) and the lme4 package (Bates *et al.* 2015). Two linear mixed models (LMM) fit by maximum likelihood (using Satterthwaite approximations to degrees of freedom to calculate p -values) were generated for the acoustic measures (one for f_0 direction and one for duration). In both LMMs, ‘Finality category’ (three dummy-coded levels: ff, nf, nn) was included as fixed factor and ‘Speaker’ was included as random intercept (the maximal converging model). Note that the incorporation of speaker as random intercept not entirely accounts for the violation of the independence assumption (Winter & Grice 2021) as multiple observations (words) were taken for each speaker (see also Table A1 in the appendix). Both models were then analysed in a post-hoc analysis in order to pairwise compare each finality category using the Tukey honestly significant difference (HSD) test.

2.5. Results

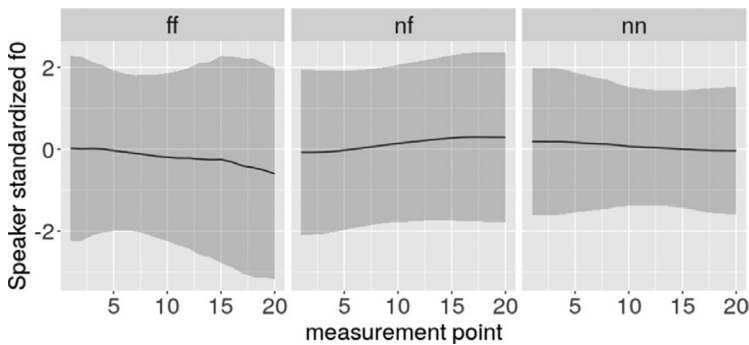
The results of the acoustic measurements (Table 5, Table 6, and Figure 5) show that syllables in phrase-final and word-final position (ff) are on average produced with a falling f_0 movement and significantly lengthened (approximately 2.7 times) compared to syllables in the other finality categories. Phrase-non-final syllables in word-final position (nf) show on average a rising f_0 movement, whereas phrase-non-final syllables in word-non-final positions (nn) show a slight falling movement. The f_0 direction measure

Table 5. Results of the acoustic measurements of f_0 direction (standardized) and duration (ms) in each finality category (ϕ_w), with standard deviations between brackets.

Measurement	ff	nf	nn
f_0 direction	-0.62 (1.39)	0.37 (0.79)	-0.22 (0.82)
Duration	318.76 (241.63)	120.24 (41.64)	118.24 (40.16)

Table 6. Pairwise comparisons for each finality category (ϕw) on the LMMs for f_0 direction (standardized) and duration (ms).

Pairwise comparison	f0 Direction				Duration			
	β	SE	z	p	β	SE	z	p
nf-ff	1.04	0.13	7.90	< 0.001	-214.22	3.77	-56.90	< 0.001
nn-ff	0.42	0.13	3.15	< 0.01	-209.77	3.74	-56.05	< 0.001
nn-nf	-0.62	0.13	-4.88	< 0.001	4.46	3.57	1.25	n.s.

**Figure 5.** Mean f_0 contour (speaker standardised) and standard deviation in each finality category (ff: ϕ -final w -final, nf: ϕ -non-final w -final, nn: ϕ -non-final w -non-final) as measured by the time-series f_0 analysis (20 measurement points).

was significantly different between each finality category. The duration of the phrase-final and word-final (ff) syllables was significantly longer compared to each of the other finality categories. The phrase-non final syllables (nf and nn) did not differ significantly in duration.

3. Contour clustering

The results discussed above provide a first coarse-grained analysis of the acoustic realisation of the syllables in each finality category. To obtain a more detailed overview of the variation in the f_0 movements, the time-series measurements were analysed using cluster analysis, following the method proposed in Kaland 2021. This approach consists of two components; measurement of f_0 and clustering. As for the measurements, the ones described in Section 2.3 were taken as the input for the cluster analysis. Cluster analysis essentially divides the f_0 contours into maximally different groups based on their numerical (shape) differences. The division in clusters is based on distances computed for each possible pair of observations (distance matrix). The distance metric used in the current analysis is Euclidean distance, a generic method used in previous literature (e.g. Calhoun & Schweitzer 2012, Cole & Steffman 2021). Clusters were merged in a bottom-up hierarchical fashion, such that at the initial state of the analysis all observations formed a cluster on their own (maximum amount of clusters). In subsequent steps the clusters with the smallest distance were merged, keeping the distances between the clusters maximal (complete linkage). The merging stops when all observations are in the same cluster, which is equal to a state in which nothing is clustered. The merging process follows a tree-structure (dendrogram) in which the number of clusters (X) is determined by the height in the tree, that is, the height at which an X amount of branches is found. Finding the number of clusters that best describes the data is crucial; see also Kaland 2021 for a discussion and details below on how this was done in the current study. It is crucial to note that with contour clustering, there is no direct correspondence between the number of clusters and the possible inventory of phonological units that underlie an f_0 contour in the language under

investigation. The dataset, the domain of the contour (syllable, word, phrase) and potentially many other factors affect the contour shapes and hence the clustering output. It is often informative to run multiple analyses with varying numbers of clusters, to see how observations are split or merged depending on whether one increases or decreases the number of clusters respectively. This was done in the current study and described in further detail below.

Before clustering, the data was subset based on outlying contours as detected by the contour clustering tool (Kaland 2021). Outlying contours mainly consisted of ones for which interpolation was challenging, either due to too many missing *f0* values or inaccurate estimation of *f0* in voiceless parts. Note that these missing values did not affect the calculation of the *f0* direction in Section 2.3. The outlying contours were removed from the data to obtain a more homogeneous set that would result in clusters of more comparable sizes, specifically, clusterings for which the mean contour of a certain cluster indeed represents the contours in that cluster. Although this procedure made the dataset more suitable for cluster analysis, it could not prevent contour shape differences due to local microprosodic or segmental influences (i.e. creaky phonation or consonantal effects; see discussion in Section 4). In total, 329 contours were clustered. The standardization to account for speaker differences led to contours with values centered around zero (the speaker's mean).

Several clustering rounds were carried out, starting with two clusters and going up to ten clusters. This was done because we did not know how many different contours were actually present in the data. It is important to note that the three finality categories indeed showed contour differences (Figure 5), however, there could be more (subtle) variation in the contours that could not be captured by a categorization based on the three finality categories. Furthermore, the clustering rounds provide an additional indication of the extent to which the shape differences among the clustered contours matches with their position at the word or phrase level, as distinguished by the finality categories.

How the contours were divided over the finality categories for each round of clustering can be seen in Table 7. Starting from six clusters, there appeared one cluster that contains four contours. This is cluster number 5 (round 6 and 7), cluster number 6 (round 8), and cluster number 7 (round 9 and 10) respectively. The asymmetrical outcome is one indication that the analysis might not be representative, given the low number of contours. In addition, an evaluation metric based on information costs to describe the data was applied to find the ideal number of clusters for this data set, following the method in Kaland & Ellison 2023. This metric suggested five contours as the most informative description of the contour differences (Table 7). The mean *f0* contours of the round with five clusters are plotted in Figure 6.

The contours in Figure 6 can be roughly divided into falling and rising movements. Clusters 1, 2, and 5 all show a falling movement, comparable in steepness but different in register. That is, the fall is either just below the speaker's mean (cluster 1), just above it (cluster 2), or well below it (cluster 5). The two rises differ in both steepness and register. That is, the rise just below the speaker's mean (cluster 3) is steeper than the one that is well above the speaker's mean (cluster 4). How the five contours are divided over the finality categories can be seen from Table 7. Examples of each of the five contours are given in Figures 7–11 and examples 8–12. The contour in cluster 1 (mid fall) appears approximately half the time in ff syllables (25/51). The contour in cluster 2 (high fall) appears more than half of the time in nn syllables (63/122). Almost half of the contours in cluster 3 (low rise) appear in nf syllables (42/88), although a large number of contours in that cluster also appears in nn syllables. The contours in cluster 4 (high rise) are mainly found in nf syllables (31/48). More than half of the contours in cluster 5 (low fall) are found in ff syllables (12/20).

(8) Example phrase with a mid fall grouped under cluster 1 (ff syllable: *li*).

<i>huli</i>	<i>Ana</i>	<i>ru</i>	<i>inggik</i>	<i>olog</i>	<i>tu</i>	<i>nakol</i>	<i>huli</i>
<i>huli</i>	<i>Ana</i>	<i>tu</i>	<i>inggik</i>	<i>olog</i>	<i>tu</i>	<i>n-akol</i>	<i>huli</i>
<i>girl</i>	<i>PN</i>	<i>DET</i>	<i>hand</i>	<i>sit</i>	<i>DET</i>	<i>1s.POSS-afraid</i>	<i>girl</i>

'This girl Ana, her hands are clumsy, I am afraid, that girl'

Table 7. Number of contours in each clustering round (N clusters) as divided over the finality categories (ϕw) and clusters (cluster number), including their information cost (shaded area indicates round with lowest information cost).

N Clusters	ϕw	Cluster number										Information cost
		1	2	3	4	5	6	7	8	9	10	
2	ff	37	52									2100.99
	nf	19	100									
	nn	15	106									
3	ff	37	44	8							1925.02	
	nf	19	69	31								
	nn	15	97	9								
4	ff	37	32	12	8						1823.58	
	nf	19	27	42	31							
	nn	15	63	34	9							
5	ff	25	32	12	8	12					1819.63	
	nf	13	27	42	31	6						
	nn	13	63	34	9	2						
6	ff	25	32	12	7	1	12				1835.08	
	nf	13	27	42	30	1	6					
	nn	13	63	34	7	2	2					
7	ff	25	24	12	7	1	12	8			1829.55	
	nf	13	26	42	30	1	6	1				
	nn	13	56	34	7	2	2	7				
8	ff	25	24	9	7	3	1	12	8		1831.91	
	nf	13	26	30	30	12	1	6	1			
	nn	13	56	29	7	5	2	2	7			
9	ff	25	24	9	4	3	3	1	12	8	1869.61	
	nf	13	26	30	13	17	12	1	6	1		
	nn	13	56	29	3	4	5	2	2	7		
10	ff	25	24	9	4	3	3	1	11	8	1	1927.58
	nf	13	26	30	13	17	12	1	5	1	1	
	nn	13	56	29	3	4	5	2	2	7	0	

- (9) Example phrase with a high fall grouped under cluster 2 (nn syllable: *ne*).
- an de nebe tot tu himin nuabuk tehen*
 an te n-ebe tot tu ha-in n-abuk tu-ehen
 1s TOP 1s.POSS-body small PRX see-2s.IMP 1s.POSS-meet do-2s.PRS
 'I am small, (but) see that I can strike'

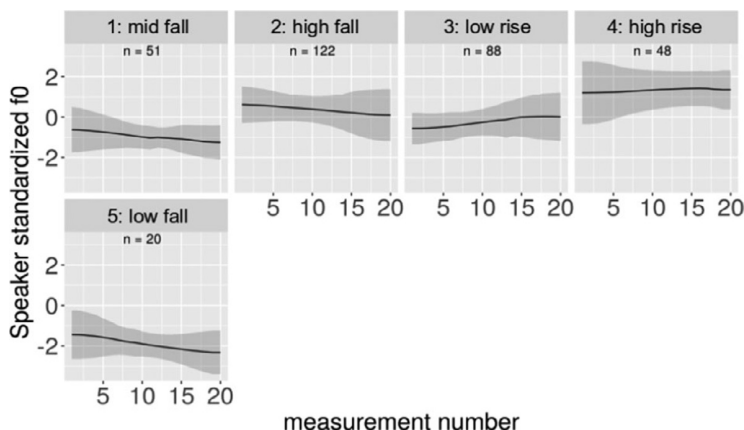


Figure 6. Mean f0 contour per cluster with five clusters assumed.

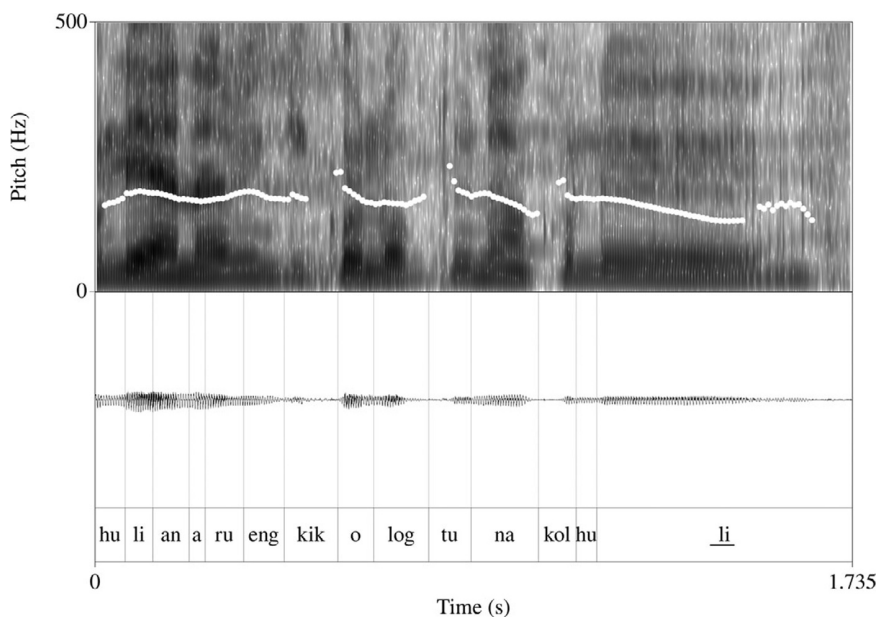


Figure 7. Example phrase with a mid fall grouped under cluster 1 (ff syllable: *li*).

- (10) Example phrase with a low rise grouped under cluster 3 (nf syllable: *go*).

hiyap itno singgonggo fam lahakekteg horiyeg
hiyap itno singgonggo fam laha-ehke-teg horiye-eg
 woman DET bike on go.up-3s.IM.PST-SS.SEQ sit-ATTR.PART
 ‘the woman is sitting on the bike’

- (11) Example phrase with a high rise grouped under cluster 4 (nf syllable: *ha* in *laha*).

oke ke laha ulug nuk ahareg suruk ari
eke ke laha ulug nuk aha-teg su-tuk ari
 and Q 3s.IM.PST:go THINK confused 3s.IM.PST:become-SS.SEQ do-PROG DEM.S
 ‘“Where did it go” he was wondering, getting confused’

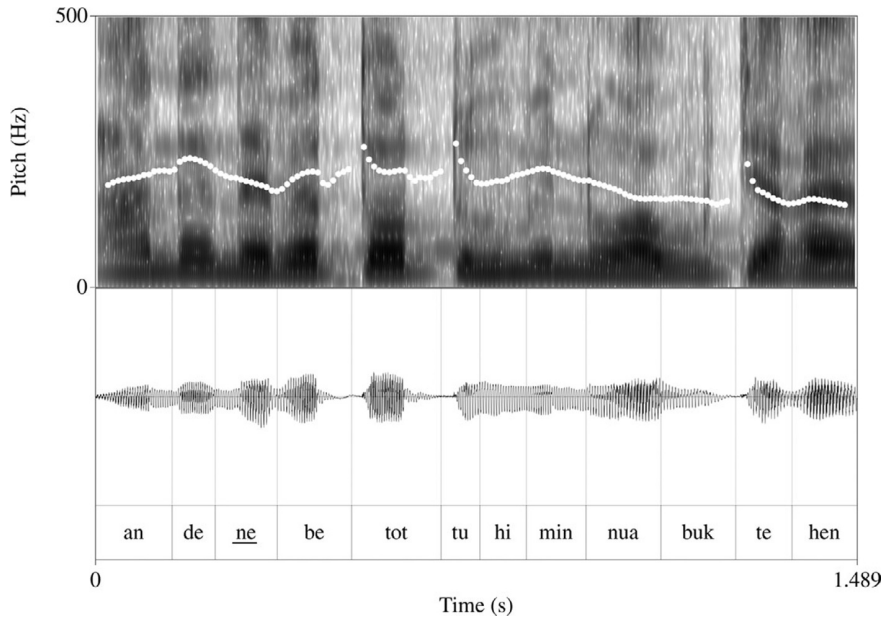


Figure 8. Example phrase with a high fall grouped under cluster 2 (nn syllable: *ne*).

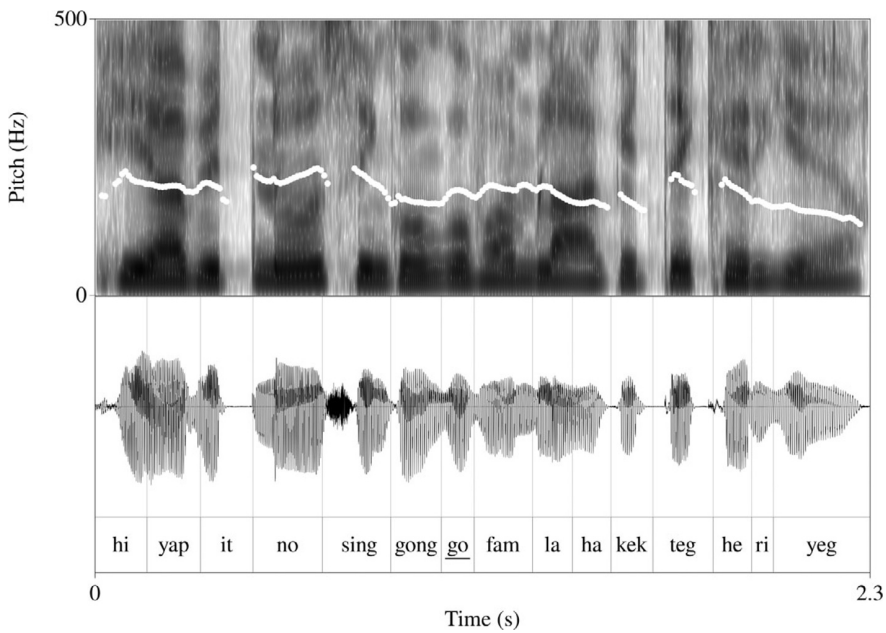


Figure 9. Example phrase with a low rise grouped under cluster 3 (nf syllable: *go*).

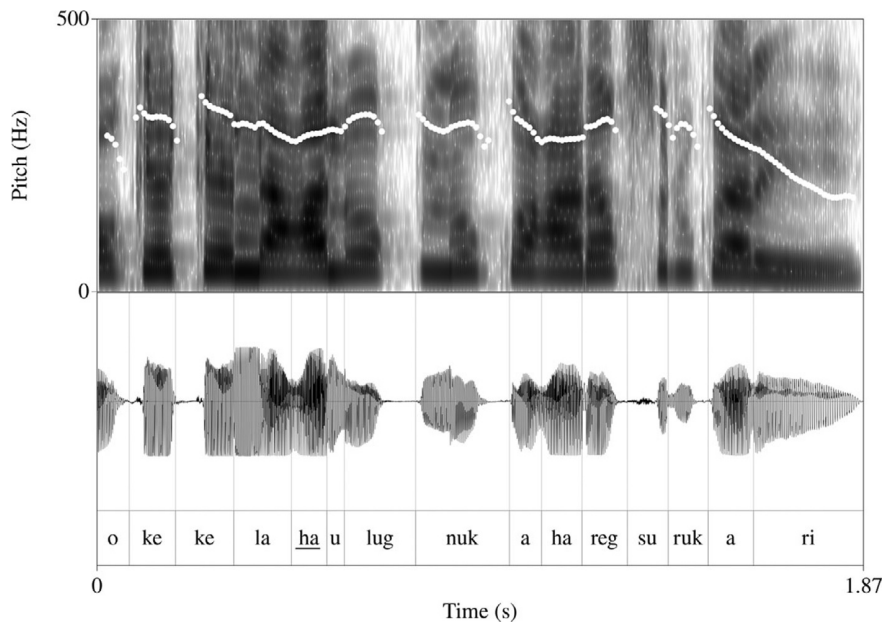


Figure 10. Example phrase with a high rise grouped under cluster 4 (nf syllable: *ha* in *laha*).

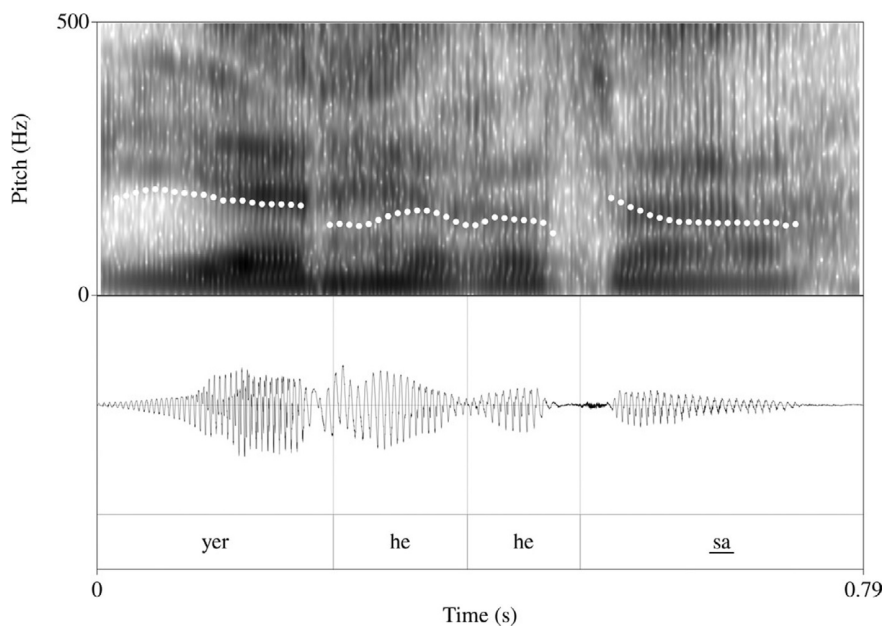


Figure 11. Example phrase with a low fall grouped under cluster 5 (ff syllable: *sa*).

(12) Example phrase with a low fall grouped under cluster 5 (ff syllable: *sa*).

yer *hehesa*
yet *ha-ehesa*
see *see-3p.IM.PST*
‘they saw’

To further clarify the relative amount of contours in each finality category and in each cluster, two additional proportion tables are presented. In Table 8 the proportions of contours in each finality category add up to 1 (horizontal), whereas in Table 9 the proportions of contours in each cluster add up to 1 (vertical).

4. Discussion and conclusion

We now turn to the two research questions (RQs) outlined in Section 1.3. Concerning RQ I, the present study provides strong support for word-final high tones. The support is found in the type of contours that are dominant in phrase-non-final syllables (nf). These are the low rise and the (shallow) high rise found in cluster 3 and 4 respectively (Table 7 and Figure 6). The low rise and high rise appear the most common contours compared to other contours found on nf syllables (Table 8) and compared to other finality categories for which these contours occur (Table 9).

It needs to be noted that there is overlap in the sense that the rises occur in other finality categories too. Moreover, other contours are found for the nf syllables. For example, low falls (cluster 5) were found for six nf syllables (Table 7). This would constitute the clearest counter-evidence for the alignment of a rise with word-final syllables in Yali. However, a post-hoc inspection of these six syllables revealed that all of them were produced with creaky voice, which could significantly have lowered the f_0 (e.g. Keating et al. 2015). It is important to acknowledge that our data comes from unscripted speech production and we cannot conclude here whether the observed differences indeed constitute noticeable ones for listeners, or whether creaky-voiced productions constitute a phonologically different category.

Table 8. Proportions of contours in each finality category and in each cluster. Proportions add up to 1 for each finality category (rows).

	1 Mid fall	2 High fall	3 Low rise	4 High rise	5 Low fall
ff	0.28	0.36	0.13	0.09	0.13
nf	0.11	0.23	0.35	0.26	0.05
nn	0.11	0.52	0.28	0.07	0.02

Table 9. Proportions of contours in each finality category and in each cluster. Proportions add up to 1 for each cluster (columns).

	1 Mid fall	2 High fall	3 Low rise	4 High rise	5 Low fall
ff	0.49	0.26	0.14	0.17	0.60
nf	0.25	0.22	0.48	0.65	0.30
nn	0.25	0.52	0.39	0.19	0.10

The shallow high fall cluster (2) contains the highest number of contours compared to the other clusters, and these contours from cluster 2 were also found on twenty-seven phrase-non-final, word-final syllables, although with a low proportion. Two remarks should be made here. First, the large number of contours in cluster 2 indicate that variation can be expected among the contours in that cluster. This variation is probably too subtle to be clustered in a more informative way. Second, the mean contour of cluster 2 remains above the speaker's mean throughout the contour and it is shallower than the low fall (cluster 5). Both observations ask for a more precise look at which features of the contour shape in word-final syllables are relevant. That is, the contours found in 84% of the nf syllables (Table 8) are either a rise or a fall that is found in a high register of the speaker. This leaves two features for further scrutiny; both the rising movement and the overall target level might be important phonologically on word-final syllables. Numerically, however, the rising movement would stand out as the most salient feature as 61% of the contours on nf syllables show it.

As for RQ II, the results show a clear interference between the word and phrase levels. In phrase-final syllables (ff) contours tend to be low and/or falling (Table 8 and Table 9). The eight contours on ff syllables found with high rises (cluster 4) were further inspected as these would most clearly counter this general tendency. The inspection revealed that they were often located at the end of short phrases with a length of one or two words (6/8) and immediately followed by phrases elaborating what was said in the short phrase; see example 13. Similarly, for the twelve ff syllables with a low rise (cluster 3), eleven of them were followed by another phrase in which the speaker elaborated the same discourse topic. Generally, the phrase following a low rise started later than the phrase following a high rise. Taken together, it seems therefore that phrasal demands can override the word-final rise such that the phrase-level tone takes precedence over the word-level tone. That is, most phrases end with a falling movement, most likely signalling discourse finality, whereas the rises found in that position clearly signal that more discourse material is upcoming, either immediately (high rise) or up to 1,400 ms later (low rise). We can therefore conclude that the word-final rise in Yali is most clearly realised in non-final phrase positions.

(13) Example of *ma* (ff) with a high rise (cluster 4).

<i>Unggul uag^{ma}.</i>	<i>Horog tot</i>	<i>sue hag toho</i>	<i>ambiyang te.</i>	<i>Yat ano horog.</i>
unggul uag=mu.	horog tot	sue hag tu-oho	ambiyang te.	yat ano horog.
head bone=LOC.	little	small bird like	do-ADV.PART beak	TOP. long Q short.

'On its head. It looks like a bird and its beak. Long or short?'

The nn syllables are most often realised with a high fall (cluster 2). A falling movement would indeed be expected without phrase- or word-level demands for specific tonal targets, that is, following the natural declination of f_0 . The somewhat higher start of this contour could also be the result of interpolation. Although a fair amount of low rises (cluster 3) also occurs on nn syllables, the majority of these thirty-four cases occurred on penultimate syllables ($N = 19$) and/or were realised on words starting with [h] ($N = 23$). Both observations are indicative of a low f_0 on these syllables, either due to an anticipatory movement toward the final rise or a lowering effect due to voiced obstruents (Kingston 2007). The [h] lowering can be furthermore observed from (subtle) f_0 dips in example 4.

The results of this study are a first exploration of how word-level and phrase-level prosody are realised and how they might interact. Two important phonological aspects remain to be investigated in more detail. One aspect pertains to the status of the rising f_0 movement in the prosodic structure. From example 5 it appears that the final position of the rise is determined after morphological operations such as cliticization and/or affixation. This would suggest that the f_0 movement is applied post-lexically. Whether this level is the one of the phonological word or clitic group (Nespor & Vogel 2007) remains to be seen. Another possibility would be to analyse the rise as a marker of the AP, which coincides in most cases with the word domain or is slightly larger, as in Korean (Jun 1996). It is furthermore likely that the occurrence of the final rise is affected by multiple factors that were not directly investigated in the current

study. From the examples, it seems that the final rise does not appear in all cases but rather shows an overall trend (Table 8 and Table 9). To conclude which (phonological) factors could underlie the realisation of the final rise, more research is needed on the phonological and prosodic structure of Yali. Second, it seems unlikely that Yali has stressed syllables, where stress would then be signaled at the word level by other cues than f_0 . The duration measures in this study provide a first coarse indication that there are no stress differences between final and non-final syllables at the word level (i.e. both in non-final phrase positions; nf, nn). Although duration is crosslinguistically a reliable cue to word stress (Gordon & Roettger 2017), and would be expected to be greater in stressed syllables than in unstressed ones, the only effect of duration found in this study could be ascribed to phrase-final lengthening. An investigation of other stress cues such as vowel quality or spectral tilt could shed more light on this issue.

Concerning the two prosodic typologies discussed in Section 1.1, it seems that Yali can be tentatively categorised as a non-stress-accent language with only post-lexical pitch (as Bengali; Table 1). In Gordon's typology (Table 2) Yali would most likely be classified as a language without edge repulsion at the word level due to the word-final rise. As for macro-rhythm (Table 3), Yali would most plausibly fit in the category of languages that have it to a strong degree due to the marking of word ends with rises. It is therefore plausible that Yali can be classified as a head/edge or edge language. The latter classification would fit an analysis in which word-final rises are post-lexically applied. Whether or not Yali makes use of pitch accents at levels higher than the phonological word remains to be investigated. This study found clear support for the use of phrase-final boundary tones (falls), but does not provide evidence on the use of pitch accents at the phrase level. Depending on future research on this issue, Yali would be classified as either a symmetrical language (no edge repulsion at word and phrase level; like Hebrew and Farsi) or as an asymmetrical language (no edge repulsion at the word level and no pitch accents; like Wolof).

It is challenging to compare the current results with the ones available for the other Dani languages (see Section 1.2: 'Dani languages: summary'), in particular given the methodological differences between the studies. Previous work has mainly provided impressionistic descriptions, whereas the current study is the first to provide an in-depth acoustic analysis. While the generalization in terms of a word-final prosodic event is shared between the previous and current results, the only acoustic evidence for Yali so far concerns f_0 . It is therefore more plausible that Yali has a word-level (pitch) accent system rather than word stress. This observation could cast doubt on the claims that attribute high f_0 to word stress, as done in the previous literature for the other Dani languages. The Yali word-level accent's identity is notably subject to or part of a phrase-level interaction, along the lines of what has been suggested for Lower Grand Valley Dani (Bromley 1961 and Section 1.2).

To conclude, this study has shown that Apahapsili Yali makes use of word-final rising f_0 contours on content words and that these can be overwritten by phrase-level boundary tones. Although this study has strongly hinted at post-lexical levels, the exact phonological level at which f_0 requirements are applied remains to be investigated. The minimal assumption based on the current outcomes constitutes two levels in the Yali prosodic structure; the phonological word for the word-final rises and the (intermediate/intonational) phrase for boundary tones. Related, more detailed investigations on the presence of pitch accents are needed to obtain a more complete profile of post-lexical prosody in Yali. Likewise, it remains a question for future research whether word stress is indeed absent in this language, as tentatively concluded in this study. Although several issues just mentioned remain unanswered, this study has shown that an initial exploration of f_0 contours in words using acoustic measurements and cluster analysis provides a methodologically novel and representative account of the type of prosodic variation found in (semi-)spontaneous Yali.

Acknowledgements. The authors thank Sonja Riesberg for comments on an earlier version of this paper, and two anonymous reviewers for their valuable feedback.

Data Availability Statement. Data collection information is provided in Riesberg et al. (2012).

Conflict of interest statement. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding disclosure statement. The research for this paper has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 281511265—SFB 1252 Prominence in Language.

Ethics statement. Informed consent was given by all speakers in the corpus that was used for the research conducted for this paper.

References

- Anderbeck, Karl. 2015. Portraits of language vitality in the languages of Indonesia. *Language documentation and cultural practices in the Austronesian world* (Papers from ICAL 12 (A–PL 019), vol. 4), ed. by I. Wayan Arka, Ni Luh Nyoman Seri Malini, and Ida Ayu Made Puspani, 19–47. Canberra: The Australian National University.
- Barclay, Peter. 2008. *A grammar of Western Dani*. (LINCOP grammar handbooks 1.) Munich: LINCOP Europa.
- Bates, Douglas; Martin Mächler; Ben Bolker; and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).
- Beckman, Mary E. 1986. *Stress and non-stress accent*. (Netherlands phonetic archives 7.) Dordrecht: Foris. OCLC: 246553005.
- Beckman, Mary E., and Jennifer J. Venditti. 2010. Tone and intonation. *The handbook of phonetic sciences*, 1st edn., ed. by William J. Hardcastle, John Laver, and Fiona E. Gibbon, 603–52. Oxford: Wiley. DOI: [10.1002/9781444317251.ch16](https://doi.org/10.1002/9781444317251.ch16).
- Boersma, Paul, and David Weenink. 2019. Praat: Doing Phonetics by Computer. Online: <http://www.praat.org/>.
- Bromley, H. Myron. 1961. *The phonology of Lower Grand Valley Dani*. Dordrecht: Springer Netherlands. DOI: [10.1007/978-94-017-5916-8](https://doi.org/10.1007/978-94-017-5916-8).
- Bruce, Gösta. 1977. *Swedish word accents in sentence perspective*. (Travaux de l'Institut de linguistique de Lund 12.) Lund: LiberLäromedel/Gleerup.
- Burung, Willem. 2007. *The phonology of Wano*. (SIL electronic working papers 2007–003.) Dallas: SIL International. Online: <http://www.sil.org/silewp/abstract.asp?ref=2007-003>.
- Burung, Willem. 2017. *A grammar of Wano*. Oxford: University of Oxford Ph.D. dissertation.
- Calhoun, Sasha, and Antje Schweitzer. 2012. Can intonation contours be lexicalised? Implications for discourse meanings. *Prosody and meaning*, ed. by Gorka Elordieta and Pilar Prieto, 271–327. Berlin: De Gruyter. DOI: [10.1515/9783110261790.271](https://doi.org/10.1515/9783110261790.271).
- Chafe, Wallace L. 1980. *The pear stories: Cognitive, cultural and linguistic aspects of narrative production*. Norwood, NJ: Praeger.
- Chafe, Wallace L. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Cole, Jennifer, and Jeremy Steffman. 2021. The primacy of the rising/non-rising dichotomy in American English intonational tunes. *Proceedings of the 1st International Conference on Tone and Intonation (TAI)*, Sonderborg, 122–26. DOI: [10.21437/TAI.2021-25](https://doi.org/10.21437/TAI.2021-25).
- Crompton, Andrew S. 1980. Intonation: Tonic stress marks versus levels versus configurations. *WORD* 31(2). 151–98. DOI: [10.1080/00437956.1980.11435689](https://doi.org/10.1080/00437956.1980.11435689).
- Eberhard, David M.; Gary F. Simons; and Charles D. Fennig. 2020. *Ethnologue: Languages of the world*. 23rd edn. Dallas: SIL International. Online: <http://www.ethnologue.com>.
- Etherington, Paul. 2002. *Nggem morphology and syntax*. Darwin: Charles Darwin University MA thesis.
- Fahner, Christiaan. 1979. *The morphology of Yali and Dani: A descriptive and comparative analysis*. Leiden: Rijksuniversiteit Ph.D. dissertation.
- Gordon, Matthew K. 2014. Disentangling stress and pitch-accent: A typology of prominence at different prosodic levels. *Word Stress*, ed. by Harry van der Hulst, 83–118. Cambridge: Cambridge University Press. DOI: [10.1017/CBO9781139600408.005](https://doi.org/10.1017/CBO9781139600408.005).
- Gordon, Matthew K., and Timo Roettger. 2017. Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard* 3(1):20170007. DOI: [10.1515/lingvan-2017-0007](https://doi.org/10.1515/lingvan-2017-0007).
- Grimes, Barbara F. (ed.) 1984. *Languages of the world: Ethnologue*. 10th edn. Dallas: Wycliffe Bible Translators.
- Hammarström, Harald; Robert Forkel; and Martin Haspelmath (eds.) 2021. *Glottolog 4.4.7*. Jena: Max Planck Institute for the Science of Human History. Online: <https://glottolog.org/>.
- Hayes, Bruce, and Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language and Linguistic Theory* 9(1). 47–96. DOI: [10.1007/BF00133326](https://doi.org/10.1007/BF00133326).
- Heeschen, V. 1998. *An ethnographic grammar of the Eipo language spoken in the central mountains of Irian Jaya (West New Guinea), Indonesia* (Mensch, Kultur und Umwelt im zentralen Bergland von West-Neuguinea, 23. Beitrag). Berlin: D. Reimer. OCLC: ocm40934802.
- Himmelman, Nikolaus P., and D. Robert Ladd. 2008. Prosodic description: An Introduction for fieldworkers. *Language Documentation & Conservation* 2(2). 244–74. Online: <http://scholarspace.manoa.hawaii.edu/handle/10125/4345>.
- Himmelman, Nikolaus P.; Meytal Sandler; Jan Strunk; and Volker Unterladstetter. 2018. On the universality of intonational phrases: A cross-linguistic interrater study. *Phonology* 35(2). 207–45. DOI: [10.1017/S0952675718000039](https://doi.org/10.1017/S0952675718000039).
- Hyman, Larry. 2006. Word-prosodic typology. *Phonology* 23(2). 225–57. DOI: [10.1017/S0952675706000893](https://doi.org/10.1017/S0952675706000893).
- Jun, Sun-Ah. 1996. *The phonetics and phonology of Korean prosody: Intonational phonology and prosodic structure*. Columbus: The Ohio State University Ph.D. dissertation.

- Jun, Sun-Ah (ed.) 2005. *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Jun, Sun-Ah (ed.) 2014. *Prosodic typology II: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Kaland, Constantijn. 2021. Contour clustering: A field-data-driven approach for documenting and analysing prototypical f0 contours. *Journal of the International Phonetic Association* 53(1). 159–88. <https://doi.org/10.1017/S0025100321000049>.
- Kaland, Constantijn, and T. Mark Ellison. 2023. Evaluating cluster analysis on f0 contours: An information theoretic approach on three languages. *Proceedings of the 20th International Congress of Phonetic Sciences*, ed. by Radek Skarnitzl and Jan Volín, 3448–52. Prague: Guarant International.
- Keating, Patricia A.; Marc Garellek; and Jody Kreiman. 2015. Acoustic properties of different kinds of creaky voice. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow. Online: <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0821.pdf>.
- Kingston, John. 2007. Segmental influences on F0: Automatic or controlled? *Tones and tunes, vol. 2: Experimental studies in word and sentence prosody*, ed. by Aditi Lahiri, Carlos Gussenhoven, and Tomas Riad, 171–210. Berlin: Mouton de Gruyter. DOI: [10.1515/9783110207576.2.171](https://doi.org/10.1515/9783110207576.2.171).
- Ladd, D. Robert. 2008. *Intonational phonology*. (Cambridge studies in linguistics.) 2nd edn. Cambridge: Cambridge University Press.
- Lekeney, Jeanete G. 2014. Causative construction in Biak, Papuan Malay, and Yali: A typological study. Manokwari: Universitas Negeri Papua MA thesis.
- Lewis, M. Paul; Gary F. Simons; and Charles D. Fennig. 2015. *Ethnologue: Languages of Asia*. Dallas: Summer Institute Linguistics.
- Lewis, Melvyn, and Gary Simons. 2010. Assessing endangerment: Expanding fishman's GIDS. *Revue Roumaine de Linguistique* 55(2). 103–20. Online: <https://www.lingv.ro/RRL%202%202010%20art01Lewis.pdf>.
- Lieberman, Philip. 1965. On the acoustic basis of the perception of intonation by linguists. *WORD* 21(1). 40–54. DOI: [10.1080/00437956.1965.11435417](https://doi.org/10.1080/00437956.1965.11435417).
- Lindström, Eva, and Bert Remijsen. 2005. Aspects of the prosody of Kuot, a language where intonation ignores stress. *Linguistics* 43(4). 839–70. DOI: [10.1515/ling.2005.43.4.839](https://doi.org/10.1515/ling.2005.43.4.839).
- Mayer, Mercer. 1969. *Frog, where are you?* New York: Dial Books for Young Readers.
- Mixdorff, Hansjörg, and Oliver Niebuhr. 2013. The influence of F0 contour continuity on prominence perception. *Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech 2013)*, Lyon, 230–34. DOI: [10.21437/Interspeech.2013-73](https://doi.org/10.21437/Interspeech.2013-73).
- Nespor, Marina, and Irene Vogel. 2007. *Prosodic phonology: With a new foreword*. New York: Mouton de Gruyter.
- Ota, Mitsuhiko; Naoto Yamane; and Reiko Mazuka. 2018. The effects of lexical pitch accent on infant word recognition in Japanese. *Frontiers in Psychology* 8:2354. DOI: <https://doi.org/10.3389/fpsyg.2017.02354>.
- Palmer, Bill (ed.) 2018. *The languages and linguistics of the New Guinea area: A comprehensive guide*. (The world of linguistics, vol. 4.) Berlin: De Gruyter Mouton.
- Pierrehumbert, Janet, and Mary E. Beckman. 1988. *Japanese tone structure*. (Linguistic inquiry monographs 15.) Cambridge, MA: MIT Press.
- Pierrehumbert, Janet, and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. *Intentions in communication*, ed. by Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, 271–311. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7916/d8kd24fp>.
- Pike, Kenneth L. 1945. *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Purba, Theodore; Onesimus Warwer; and Reimundus Fatubun. 1993. *Fonologi bahasa Dani Barat*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.
- R Core Team. 2020. R: the R project for statistical computing. Online: <https://www.r-project.org/>.
- R Studio Team. 2020. RStudio: Integrated development for R. Online: <https://www.rstudio.com/>.
- Remijsen, Bert, and Vincent J. Van Heuven. 2005. Stress, tone and discourse prominence in the Curaçao dialect of Papiamentu. *Phonology* 22(2). 205–35. DOI: [10.1017/S0952675705000540](https://doi.org/10.1017/S0952675705000540).
- Riad, Tomas. 2003. Diachrony of the Scandinavian accent typology. *Development in prosodic systems*, ed. by Paula Fikkert and Haike Jacobs, 91–144. Berlin: De Gruyter. DOI: [10.1515/9783110894530.91](https://doi.org/10.1515/9783110894530.91).
- Riesberg, Sonja. 2021. Introduction to the Yali – English – German dictionary with a short grammatical sketch. In Riesberg et al. 2021, 1–47.
- Riesberg, Sonja; Carmen Dawuda; Lucas Haiduck; Nikolaus Himmelmann; and Kurt Malcher. 2021. *A Yali (Angguruk) – English – German dictionary*. Canberra: Asia-Pacific Linguistics. Online: <http://hdl.handle.net/1885/127381>.
- Riesberg, Sonja; Kristian Walianggen; and Siegfried Zöllner. 2012. Yali. The language archive. [Online archive] Online: <https://hdl.handle.net/1839/00-0000-0000-0017-EA2F-C>.
- Rose, Phil. 1987. Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication* 6(4). 343–52. DOI: [10.1016/0167-6393\(87\)90009-4](https://doi.org/10.1016/0167-6393(87)90009-4).
- Savino, Michelina; Loredana Lapertosa; Alessandro Caffò; and Mario Refice. 2019. Prosodic convergence in Italian game dialogues. *Proceedings of the 7th Tutorial and Research Workshop on Experimental Linguistics*, Saint Petersburg, 151–54. DOI: [10.36505/ExLing-2016/07/0034/000293](https://doi.org/10.36505/ExLing-2016/07/0034/000293).

- Sawaki, Yusuf. 1997. The verbs of Middle Yali. Jayapura: Universitas Cenderawasih MA thesis.
- Senft, Gunter. 2007. The Nijmegen space games: Studying the interrelationship between language, culture and cognition. *Person, space and memory in the contemporary Pacific: Experiencing new worlds*, ed. by Jürg Wassmann and Katharina Stockhaus, 224–44. New York: Berghahn Books. Online: <https://hdl.handle.net/11858/00-001M-0000-0013-1D6A-9>.
- SIL International. 2021. Field linguist's toolbox. Online: <https://software.sil.org/toolbox/>.
- Silverman, Kim E. A., and Janet B. Pierrehumbert. 1987. The timing of prenuclear high accents in English. *The Journal of the Acoustical Society of America* 82(S1):S19. DOI: <https://doi.org/10.1121/1.2024693>.
- Silzer, Peter J., and Heljä Heikkinen-Clouse. 1991. Index of Irian Jaya languages. Special Issue of *Irian: Bulletin of Irian Jaya*. Jayapura: Program Kerjasama Universitas Cenderawasih and SIL 2nd edn.
- Skopeteas, Stavros; Ines Fiedler; Samantha Hellmuth; Anne Schwarz; Ruben Stoel; Gisbert Fanselow; Caroline Féry; and Manfred Krifka. 2006. *Questionnaire on information structure (QUIS): Reference manual*. Vol. 4. Potsdam: Universitätsverlag Potsdam.
- Usher, Timothy, and Edgar Suter. 2021. Newguineaworld. Online: <https://newguineaworld.linguistik.uzh.ch/>.
- van der Hulst, Harry. 2012. Deconstructing stress. *Lingua* 122(13). 1494–1521. [10.1016/j.lingua.2012.08.011](https://doi.org/10.1016/j.lingua.2012.08.011).
- Voorhoeve, C. L. 1975. *Languages of Irian Jaya: Checklist, preliminary classification, language maps, wordlists*. Canberra: Australian National University.
- Walianggen, Kristian. 2012a. Affixation in Yali language of Apahapsili dialect. Manokwari: Universitas Negeri Papua MA thesis.
- Walianggen, Yehemia. 2012b. Verb agreement in Yali: A Trans New Guinea language of Dani family. Manokwari: Universitas Negeri Papua MA thesis.
- Wandik, Edward. 2014. Reduplication in Yali. Manokwari: Universitas Negeri Papua MA thesis.
- Wilson, John D. 1999. *Scripture in an oral culture: The Yali of Irian Jaya*. Edinburgh: University of Edinburgh Ph.D. dissertation.
- Winter, Bodo, and Martine Grice. 2021. Independence and generalizability in linguistics. *Linguistics* 59(5). 1251–77. DOI: [10.1515/ling-2019-0049](https://doi.org/10.1515/ling-2019-0049).

Appendix

Table A1. Overview of the number of syllables per finality category (ff, nf, nn) per speaker (1–29) in the final subset of syllables analysed in this study ($N = 376$).

spk	ff	nf	nn
1	2	2	1
2	1	2	1
3	6	1	0
4	1	0	2
5	6	5	6
6	4	7	13
7	8	1	1
8	2	1	0
9	2	2	3
10	0	1	1
11	11	25	21
12	2	3	5
13	11	12	11
14	2	4	5
15	4	3	3

(Continued)

Table A1. Continued

spk	ff	nf	nn
16	0	0	1
17	1	3	0
18	8	10	8
19	5	0	1
20	2	2	1
21	4	3	4
22	10	13	8
23	7	18	18
24	6	1	5
25	2	0	2
26	14	1	5
27	2	3	0
28	1	2	0
29	1	0	0

Cite this article: Kaland, Constantijn, Katharina Gayler, Martine Grice & Nikolaus P. Himmelmann. 2026. A first exploration of word- and phrase-final f0 movements in spontaneous Yali. *Phonological Data and Analysis*. 8, e4, 1–25. <https://doi.org/10.1017/S2642182825000030>